Introducing the Joint EEG-Development Inference (JEDI) Model: A Multi-way, Data Fusion Approach for Estimating Paediatric Developmental Scores Via EEG

Introducing the Joint EEG-Development Inference (JEDI) Model: A Multi-way, Data Fusion Approach for Estimating Paediatric Developmental Scores Via EEG

Eli Kinney-Lang, Ahmed Ebied, Bonnie Auyeung, Richard F.M. Chin, and Javier Escudero *Member, IEEE*

***Abstract*—Accounting for developmental changes in children is a key consideration for adapting neurorehabilitation technologies to paediatric populations. Using well-established clinical tests and questionnaires can be resource and time intensive. With many data-driven rehabilitation approaches relying on EEG data, a means to rapidly assess and infer developmental status of children directly from these recordings could be critical. This manuscript proposes a new model for estimating classic developmental diagnostic scores by exploiting data fusion in a joint tensor-matrix decomposition of the EEG and score data. We have designated this model the Joint EEG-Development Inference (JEDI) model. The proposed model is illustrated using a common EEG task (button press) via publicly available paediatric data from pre-adolescent children. Using 3 distinct recording blocks for training, validation and testing and a 10-fold cross-validation scheme, a robust experimental design was used to evaluate the JEDI model under various conditions. Results indicate that the JEDI model can estimate the developmental scores of children while maintaining a high degree of similarity at a population level. These results highlight the JEDI model as a potential evolving tool for rapidly assessing child's development. Clinically, the proposed model could provide useful developmental information in a convenient and low resource manner.**

***Index Terms*—Multi-way analysis, tensor analysis, data fusion, EEG, CPD, PARAFAC, developmental neuroscience, child development.**

I. Introduction

ADVANCEMENTS in engineering, electrophysiology and brain modelling have underpinned new avenues of rehabilitation medicine technology in recent years, such as virtual reality/environment training [1], [2] and brain-computer interfaces (BCI) [3]–[5]. Tapping into the underlying brain dynamics and neural mechanisms, like neural plasticity, has been a corner stone for these technologies [4]–[6]. Being able to exploit the vast potential of plasticity present during development is a driving force behind expanding these rehabilitation paradigms to paediatric populations [5], [7], [8]. However, the wildly heterogeneous and variable rate of development among children, especially for children in need of neurorehabilitation paradigms, have largely barred the translation of technologies like BCI [5], [7]–[9].

Recently we proposed methodology for tackling such developmental hurdles through exploiting the natural higher-order structure of paediatric electroencephalography (EEG) data via multi-way (i.e. tensor) analysis [7], [8]. In this approach, we utilize advanced signal processing to model the complex multi-dimensional EEG as a combination of low-rank component matrices in which unique features can be identified and linked to child age [7], [8]. The outlined tensor framework provided a foundation for identifying features with respect to child age, but did not take into account other potential measures of 'development', e.g. traditionally used clincial measures of child development. Therefore, there is room to expand the tensor framework model to include information from such clinical tests.

Data fusion offers a way to integrate and analyze data from multiple sources of information through joint or coupled factorization [10]–[14]. The coupled (i.e. fused) data can provide insights which are not readily apparent from a single source of data [10]–[14]. This coupling of information has been referred to as multi-view or multi-relational [10], and arises when an observation's features consist of two or more disjoint sets or views [10]. Importantly, tensor decompositions can be understood as a special case of joint factorization (i.e. data fusion) where the data sets are homogeneous [10]. Structured data fusion (SDF) [10], [15] can offer an application agnostic, customizable data fusion framework capable of jointly decomposing various data types, including tensors and matrices. Therefore, data fusion via SDF provides a tool compatible within the tensor framework to incorporate clinical metrics of child development.

EEG data and traditional developmental scores measured through psychometric tests can intuitively be recognized as multi-modal (or multi-relational) data, where the two disjoint sets reveal complimentary information regarding child development. As such, their joint tensor-matrix decomposition and SDF [10], [15] could provide novel insights into of the underlying developmental information and provide additional boons in terms of inferring development status of children. Therefore, we have termed the SDF of these data types the Joint EEG-Development Inference (JEDI) model. To the authors' knowledge, no such modelling framework has

E. Kinney-Lang, A. Ebied and J. Escudero are with the School of Engineering, Institute for Digital Communications (IDCOM), University of Edinburgh, Edinburgh, United Kingdom, EH9 3FG. Bonnie Auyeung is with the School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom EH8 9JZ. Richard Chin is affiliated with the Royal Hospital for Sick Children, Edinburgh, the Muir Maxwell Epilepsy Centre and the Centre for Clinical Brain Sciences at the University of Edinburgh, Edinburgh.

Primary author e-mail: e.kinney-lang@ed.ac.uk

Manuscript submitted June 30, 2018.

been previously described in data fusion literature. In this manuscript we:

- Outline and describe the structure for the proposed JEDI model
- Apply the JEDI model using a common EEG task (left/right button-press) and relevant clinical behavioural tests (e.g. the Wechsler Abbreviated Scale of Intelligence-Second edition)
- Demonstrate how the JEDI model can exploit shared information in order to predict developmental scores for new/unseen children based on EEG recordings

Effective modelling of the multi-relational data associated with child development supports the proposed JEDI model as a potentially useful tool for characterizing children's developmental status in emerging rehabilitation technologies.

II. Materials & Methods

A. Notation and Definitions

This paper follows the tensor notation and definitions outlined in [8], [16]. The relevant notation is summarized presently.

Tensors are a multi-way array of data, designated by a calligraphic upper-case letter (e.g. $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$). Bold upper-case letters represent matrices, which are considered as 2-way tensors (e.g. $\mathbf{X} = [a_1, a_2, \ldots, a_J] \in \mathbb{R}^{I \times J}$). Bold lower-case letters are used for vectors, i.e. 1-way tensors (e.g. $\mathbf{a}_j$). The n-th way (also referred to as domain) matrix of a given tensor $\mathbf{A}^{(n)}$ is the corresponding domain-n slice of that tensor. As an example, in a three-way tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ with elements x_{ijk}, the domain-1 frontal slice of $\mathcal{X}$ is $\mathbf{X}^{(1)} \in \mathbb{R}^{I \times JK}$, the domain-2 lateral slice is $\mathbf{X}^{(2)} \in \mathbb{R}^{J \times IK}$ and domain-3 horizontal slice is $\mathbf{X}^{(3)} \in \mathbb{R}^{K \times IJ}$ [8].

The Kronecker product of two matrices $\mathbf{A} \in \mathbb{R}^{I \times J}$ and $\mathbf{B} \in \mathbb{R}^{K \times L}$ is denoted by $\mathbf{A} \otimes \mathbf{B}$ [16]. The result is a matrix $\mathbf{T} \in \mathbb{R}^{IK \times JL}$ with:

$$\mathbf{T} = [a_1 \otimes b_1 \quad a_1 \otimes b_2 \quad a_1 \otimes b_3 \cdots a_J \otimes b_{L-1} \quad a_J \otimes b_L]$$

The Khatri-Rao product can be viewed as a 'column-wise' Kronecker product. Given two matrices $\mathbf{A} \in \mathbb{R}^{I \times K}$ and $\mathbf{B} \in \mathbb{R}^{J \times K}$, their Khatri-Rao product is $\mathbf{A} \odot \mathbf{B}$ [16], with result matrix $\mathbf{T} \in \mathbb{R}^{IJ \times K}$ [16]. This is equivalent to:

$$\mathbf{A} \odot \mathbf{B} = [a_1 \otimes b_1 \quad a_2 \otimes b_2 \quad \cdots \quad a_K \otimes b_K]$$

The outer product of vectors $\mathbf{a}$ and $\mathbf{b}$ is represented by $\mathbf{a} \circ \mathbf{b}$. We will denote the non-negative least squares regression between a matrix $\mathbf{A}$, and vector $\mathbf{d}$ by $\mathrm{NN}(\mathbf{Ad})$ [8].

B. Datasets

Data for this study was taken from a publicly available dataset provided by the Child Mind Institute (CMI) [17]. Our study focused on pre-adolescent children (< 11 y.o.), resulting in a sample size of $n = 45$ children. EEG data was captured using a high-density 129-electrode hydro-cel EEG cap, with the pre-processing handled by the CMI prior to the point of data distribution (for details on EEG pre-processing see [17]). Through utilizing the already pre-processed EEG offered by the CMI, we aimed to promote easier replication of the methodology outlined in this manuscript.

EEG data was recorded while children participated in different activities, including both resting state and task-oriented paradigms [17]. The *Contrast Change Paradigm* (CCP) task was the focus for the presented analysis. In brief, the CCP task instructs the child to pay attention to two objects on screen, each originally at 50% contrast. Randomly, over the course of a trial, one object will decrease to 0% contrast, while the other increases to 100%. The child then presses the corresponding (left/right) button indicating which object is increasing in contrast [17]. The CCP task allows isolation of specific brain processing phenomena [17], including a post-decision processing but pre-motor time window (i.e., the window after the subject has decided on an object, up to the actual button press). This pre-motor isolated brain signal was chosen as it could be considered similar in design to the pre-motor imagined movement common in rehabilitation technologies, like BCI [16]. Furthermore, the CCP task required engagement from the child, thus providing brain activity in which developmental changes could be actively investigated. A random set of 24 trials were run (12 left, 12 right) for a single block of the CCP, and each child attempted to complete 3 CCP blocks. For full details, please see [17].

Children were excluded from this study based on two main factors. First, the experimental set-up required each CCP block 1, 2 and 3 to consist of exactly the same children across all blocks in order to facilitate comparison, validation and testing. Therefore if any child was unable to complete all 3 CCP blocks they were excluded. Additionally, all children in the analysis needed to have the same developmental tests recorded. Children without the valid recorded developmental scores outlined below were also excluded. This resulted in $n = 22$ children, age 8-11 years old to be included in this study.

C. Developmental Tests

Phenotypic data was recorded alongside the paediatric EEGs. Several developmental tests were recorded for each subject. From the available developmental measurements, we chose three psychological tests to investigate. These were: 1.) the Wechsler Abbreviated Scale of Intelligence- Second Edition (WASI-II) [18]; 2.) the Wechsler Individual Achievement Test- Second Edition-Abbreviated (WIAT-IIA) [19]; and 3.) the Kid-KINDL (KK) questionnaire [20], [21]. This set of CMI phenotypic data measurements included two independent, external measures of development conducted by an expert (WASI-II, WIAT-IIA), and one self-reflective measure of development (KK). Through these three measures we aimed to capture a well-rounded set of classic 'developmental metrics'.

To narrow the scope of these developmental metrics, we *a priori* opted to focus on the phenotypic developmental scores associated with IQ (i.e. instead of behaviour, social, etc.). Thus, our analysis included the full-scale IQ (FSIQ-4) measurement from WASI-II, the full composite score for WIAT-IIA, and only the School Subscale in the KK questionnaire. The score values for WASI-II and WIAT-IIA were reported on

a normative 100 point curve (where 'normal' is considered at $\mu = 100$, $s.d. \pm 10$). The KK questionnaire, however, ranged from 0-5 points for the School Subscale based on responses 'never', 'seldom', 'sometimes', 'often' and 'all of the time', respectively. The School Subscale scores were arranged by the CMI dataset so that higher points indicated the child feeling more successful/positive in the school setting. To more easily visualize across data, the KK questionnaire's 5 point scale was translated to a 0-100 point scale for analysis.

The scaled scores from the WASI-II, WIAT-IIA and KK questionnaire were aggregated together across all children to construct a psychometric score matrix, $\mathbf{S} \in \mathbb{R}^{N \times T}$. Here, the $[Subject]$ and $[Score]$ domains are of size N, T respectively. Prior to cross-validation analysis, the score matrix was standardized to a z-score using MATLAB so that the appropriate scale of each of the scores could be recovered in the prediction step.

D. Data Processing

EEG data was originally recorded and processed as described in the CMI dataset [17]. The high-density 129-electrode EEG was recorded at a sampling rate of 500 Hz, with a bandpass of 0.1 to 100 Hz, and a reference point of Cz for each paradigm. For this paper, the desired CCP EEG recordings were then re-processed using the toolbox Fieldtrip [22] in MATLAB to find the time-frequency response (TFR). First, the data was re-referenced to a common average reference (CAR), and bandpass filtered between 5 to 18 Hz. This frequency range was selected for two purposes. First, it captures the variable range of EEG frequencies in children associated with executing motor movement in the button-press task, e.g. the shifts in frequency and power of EEG bands alpha, mu and beta which are associated with development throughout childhood [23]–[26]. The lower beta range is included as it continues to change and mature with the development of the child [24]–[27] and is associated with motor activity [25]–[27]. The upper beta range was excluded, however, to avoid potential contamination from the steady-state visual evoked signal at 20 and 25 Hz [17] present in the CCP task. A spatial filter was applied to focus on 20 central channels over the motor cortex (E103-106, E110-112, E116-118 for the right, and E7, E13, E20, E28-30, E34-36, E41 for the left). The TFR was then calculated using Fieldtrip's multitaper convolution function ('mtmconvol'), across Hanning tapers on the time series trial data [22]. The data was time-locked to the button press response, designated as time $t = 0$. Sliding time windows of 20 ms were used to estimate the TFR for two second trials defined from -1500 to 500 ms, with respect to the button press and spectrum values from the TFR were recorded. A baseline from -1500 to -1000 ms was used to calculate the relative change in the TFR spectrum data. Analysis was done separately for left/right button press data.

E. Tensor Analysis

Figure 1-A illustrates a general example of a three-way tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times N}$ corresponding to $[Spatial]$, $[Spectral]$ and $[Subject]$ domains respectively.

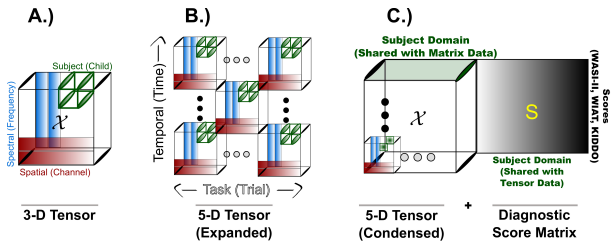

Fig. 1. Visual representation of various tensor structures. A.) A simple representation of a 3-domain tensor $\mathcal{X}$ with $[Spatial]$ (channel), $[Spectral]$ (frequency) and $[Subject]$ (child) domains. B.) Illustration of how the 3-way tensor from (A) can be stacked along $[Temporal]$ (filled dot) or $[Task]$ (open dot) modes to create a 5-domain tensor. C.) An illustration of the shared $[Subject]$ domain between a 5-way tensor $\mathcal{X}$ and a matrix $\mathbf{S}$ (e.g. developmental score matrix).

1) Constructing a Tensor: From the processed EEG data, a 5-way tensor was constructed with domains of $[Trial] \times [Channel] \times [Frequency] \times [Time] \times [Subject]$ using the TFR power spectrum data calculated for each child, at each trial and every CCP block. This 5-way tensor is visually represented in Figure 1-B. A subset of the 5-way tensor was identified to be core for defining the tensor-aspect of the joint model. The model did not need resolution at the $[Trial]$ domain level as the child's developmental scores did not change from trial to trial. Additionally, the model training and validation was done only on specific combinations of the CCP recording blocks. Thus, a sub-tensor which averaged across the whole $[Trial]$ domain and included only the relevant blocks was defined for experimental analysis. Therefore, analysis was conducted on the 4-way sub-tensor $\mathcal{X}_{sub} \in \mathbb{R}^{I \times J \times K \times N}$ corresponding to $[Channel] \times [Frequency] \times [Time] \times [Subject]$ domains, respectively, for each CCP block. Using the sub-tensor provided an additional benefit of reducing computational expenses.

2) Tensor Decomposition: The constructed 4-way tensor was decomposed via the toolbox Tensorlab [15], [28], [29] for MATLAB through Parallel Factor Analysis (PARAFAC), also commonly referred to as Canonical Polyadic Decomposition (CPD). CPD exploits the inherent higher-order structure in the EEG data [30]–[32] to define underlying latent structures in the data, and is a common decomposition model choice for EEG [16], [31]. Additionally, CPD is capable of identifying developmental profiles in paediatric EEG [7], [8].

The CPD model decomposes a multi-way tensor $\mathcal{X}$ into a (minimal) sum of rank-1 tensors [16], [33], [34]. The CPD can alternatively be considered as a special Tucker decomposition which has a super-diagonal core [16], [33], [34]. This provides a guaranteed 1:1 interaction in the identified latent factors across the modes of $\mathcal{X}$ [8], [34], [35]. This in turn helps to provide insight into the structural relationships within the data, including the ability to retain information regarding the latent developmental relationships between children at different developmental points and their respective EEG [8].

For an explicit example of the CPD decomposition, let a fourth-order tensor be $\mathcal{X} \in \mathbb{R}^{I \times J \times K \times N}$ [8]. Then a general rank-*R* CPD model is defined as

$$\mathcal{X} \approx \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \circ \mathbf{d}_r \tag{1}$$

for $r = 1, 2, .., R$ with $\mathbf{a}_r \in \mathbb{R}^I$, $\mathbf{b}_r \in \mathbb{R}^J$,$\mathbf{c}_r \in \mathbb{R}^K$ and $\mathbf{d}_r \in \mathbb{R}^N$. This can alternatively be written as

$$x_{ijkn} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} d_{nr} + e_{ijkn} \tag{2}$$

for $i = 1, ..., I; \quad j = 1, ..., J; \quad k = 1, ..., K; \quad n = 1, ..., N$ and $r = 1, ..., R$ with $x_{ijkn}, a_{ir}, b_{jr}, c_{kr}, d_{nr}$ and e_{ijkn} as elements of $\mathcal{X}$, domains $\mathbf{A} \in \mathbb{R}^{I \times R}$, $\mathbf{B} \in \mathbb{R}^{J \times R}$, $\mathbf{C} \in \mathbb{R}^{K \times R}$, $\mathbf{D} \in \mathbb{R}^{N \times R}$ and residual $\mathcal{E} \in \mathbb{R}^{I \times J \times K \times N}$, respectively [8].

F. Data Fusion and Grid Search

The EEG tensor $\mathcal{X}_{sub} \in \mathbb{R}^{I \times J \times K \times N}$ and the score matrix of $\mathbf{S} \in \mathbb{R}^{N \times T}$ were jointly decomposed using the CPD option in Tensorlab's structured data fusion (SDF) [10], [15]. The score matrix $\mathbf{S}$ included $[Subject]$ and $[Score]$ domains with component factors $\mathbf{D} \in \mathbb{R}^{N \times R}$, and $\mathbf{E} \in \mathbb{R}^{T \times R}$ respectively and $\mathbf{S} \approx \mathbf{D}\mathbf{E}^T$. The shared domain $[Subject]\mathbf{D} \in \mathbb{R}^N$ was given to the SDF model as the single overlapping domain from the multi-relational set. Figure 1-C visually illustrates the shared domain in a tensor-matrix SDF. An initial estimate for the decomposed tensor was first calculated using Tensorlab's multi-linear SVD [28] while the score matrix was initialized randomly.

Non-negativity was imposed on each domain of the EEG tensor (and thus consequently on the $[Subject]$ domain of the score matrix). The non-negativity constraint improved the ability to interpret factors of the model, and grounded them within realistic bounds [8]. For example, non-negativity constraints on the power spectrum data in the $[Frequency]$ domain are justifiable as negative values for such data would be meaningless [8]. Regularization was imposed on both the EEG tensor and score matrix data. For the EEG tensor, L2 regularization helped reduce fitting to noise for the $[Channel]$, $[Frequency]$, and $[Time]$ domains. Similarly, L2 regularization was included for the matrix's $[Score]$ domain. L1 regularization was imposed on the $[Subject]$ domain for both datasets to help promote minimal over-fitting and potentially sparse responses across the shared domain.

Using the SDF structure, the degree of contribution of the EEG tensor, score matrix, L2 and L1 regularization terms were defined in terms of their relative weights. These weights can be considered as hyperparameters of the JEDI model. The objective function to minimize the SDF of $\mathcal{X}$ and $\mathbf{S}$ is thus:

$$\begin{aligned} \min_{\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\mathbf{E},R} \quad & (\lambda_1/2)||\mathcal{X} - \mathcal{M}_{CPD}(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},R)||_F^2 + \\ & (\lambda_2/2)||\mathbf{S} - \mathcal{M}_{CPD}(\mathbf{D},\mathbf{E},R)||_F^2 + \\ & (\lambda_3/2)(||vec(\mathbf{A})||_F^2 + ||vec(\mathbf{B})||_F^2 \\ & + ||vec(\mathbf{C})||_F^2 + ||vec(\mathbf{E})||_F^2) + \\ & (\lambda_4/2)||vec(\mathbf{D})||_1 \end{aligned} \tag{3}$$

where $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}, R$ are the $[Channel]$, $[Frequency]$, $[Time]$, $[Subject]$, and $[Score]$ domain component matrices and number of components, respectively; $\mathcal{M}_{CPD}$ is the joint CPD decomposition; and λ_{1-4} are the relative weights for tensor data $\mathcal{X}$, the score matrix $\mathbf{S}$, L2, and L1 regularization respectively.

In order to estimate λ_{1-4}, we employed a grid search varying λ_{1-3} while fixing λ_4 at 1. We varied $\lambda_{1,3}$ logarithmically from 0.01 to 100 in nine steps each (i.e. 0.01, 0.0316, 0.1, 0.316...). To be sure that the relative weight of neither the tensor nor the matrix was overbearing, we varied λ_2 logarithmically proportional to λ_1 such that it was at most (least) one log-step more (less) than the value for λ_1 (i.e., for $\lambda_1 = 1$, $\lambda_2 = [0.1, 10]$), over 5 steps in log-space. For each value λ_{1-4}, the SDF was calculated for one of the three CCP blocks (e.g. CCP block 1), and validated on an independent second CCP block (e.g. CCP block 2). Validation using the second CCP block (without loss of generality) provided the relative errors for each combination of parameters. This procedure thereby approximated the generalizability of each set of hyperparameters with respect to unseen EEG data, providing insight for choosing an ideal set of hyperparameters applicable across all combinations of the CCP blocks. The third CCP block not used for either defining the initial SDF model or grid search validation was used for testing (e.g. CCP block 3).

In addition to the relative weights λ_{1-4}, the grid search varied the number of components *R* for the joint CPD decomposition between 2 to 5 components. This range for the components was chosen specifically to emphasize the tensor ranks in which the JEDI model effectively balances (1) model validity measured by percent explained variance (EV); (2) model validity of the CPD using the Core Consistency Diagnostic (CORCONDIA) [8], [36], [37] metric; and (3) computational time required. In principle, CORCONDIA indicates the suitability of using CPD to model the data and is inversely related to the EV (for interested readers and a more in depth discussion, please see [8], [36], [37]). A larger number of components used for a CPD decomposition tends to decrease the CORCONDIA while increasing the EV [8], [36], [37] and the computational time [29], [38]. As such, selecting the correct number of components to decompose a dataset requires a balance between these aspects.

The grid search was ran over all permutations of the independent CCP blocks 1, 2 and 3 for model definition, validation and testing. Results were recorded and a specific model weight with *R* components was selected for analysis.

G. Characteristic Filters and Direct Projection

An advantage of the CPD decomposition is the ability to directly define the 'characteristic filters' [8], [39] which hold the common interactions between the tensor modes. Additionally, the CPD model provides a direct multi-linear model, and has previously been used extensively with EEG data [8], [16], [39]. The information held in these characteristic filters are based on the resulting CPD component matrices, and allow new incoming data to be directly projected onto the previously estimated interactions in order to estimate the component weights of any specific mode, without needing to re-run the tensor decomposition [8], [39]. In this study, we used

characteristic filters based on the $[Channel]$, $[Frequency]$, and $[Time]$ domains to directly project new children onto the $[Subject]$ domain. Using the common interactions held in those 3 domains of $\mathcal{X}_{sub}$, we can estimate a new child's $[Subject]$ component weights directly. In practice, this was done as follows.

First, let $\mathcal{X}_{sub} \in \mathbb{R}^{I \times J \times K \times N}$ be a 4-way tensor in our analysis, with modes $[Channel] \times [Frequency] \times [Time] \times [Subject]$, and $\mathbf{S} \in \mathbb{R}^{N \times T}$ be a 2-way matrix holding developmental score data, with modes $[Subject] \times [Score]$. Then, let $\mathbf{A} \in \mathbb{R}^{I \times R}, \mathbf{B} \in \mathbb{R}^{J \times R}, \mathbf{C} \in \mathbb{R}^{K \times R}, \mathbf{D} \in \mathbb{R}^{N \times R}$ and $\mathbf{E} \in \mathbb{R}^{T \times R}$ be the resulting *R*-component factor matrices of the joint CPD factorization for each mode, respectively. Also, let a new testing tensor $\mathcal{Y}$ be of similar design to $\mathcal{X}$ (i.e. have the same domain designations), with new subjects and recordings. Then weights for the new subjects in the $[Subject]$ domain of $\mathcal{Y}$ (i.e. $\hat{\mathbf{D}}$) can be determined through direct projection on the characteristic filters of $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ by:
Unfold $\mathcal{X}_{sub}$ along the $[Subject]$ domain, defining the $[Channel] \times [Frequency] \times [Time]$ characteristic filters using the Khatri-Rao product:

$$\mathbf{X}_{sub}^{(4)} = \mathbf{D}[\mathbf{C} \odot \mathbf{B} \odot \mathbf{A}]^T, \tag{4}$$

$$\mathbf{X}_{sub}^{(4)} = \mathbf{DW} \tag{5}$$

where $\mathbf{W}$ is an encoding matrix of the interactions between the $[Channel]$, $[Frequency]$, and $[Time]$ domains [8], [39]. The projection from new subjects onto these common interactions provides estimated weights for the subjects' spatial, spectral and temporal profile patterns [8], [39]. New weights from the direct projection are then found using the non-negative least-square projection ($NN(\cdot)$) between the encoding matrix, $\mathbf{W}$, and the unfolded testing tensor $\mathcal{Y}$ along the $[Subject]$ domain as done in [8], $\mathbf{Y}^{(4)}$:

$$\mathbf{Y}^{(4)} = \mathbf{D}'[\mathbf{C}' \odot \mathbf{B}' \odot \mathbf{A}']^T, \tag{6}$$

$$\hat{\mathbf{D}} = \mathrm{NN}(\mathbf{Y}^{(4)}\mathbf{W}), \tag{7}$$

where $\hat{\mathbf{D}}$ holds the estimated spatial, spectral, and temporal weights for the new subjects [8], [39]. The non-negative least squares as an alternative to the pseudo-inverse to inhibit potential outliers from near-zero data, similar to [8].

Since $\mathcal{X}_{sub}$ is jointly factorized with the developmental score matrix $\mathbf{S}$, the new estimated $\hat{\mathbf{D}}$ will be encoded with respect to the developmental scores as well. Therefore, to estimate a new child's estimated developmental score for WASI-II, WIAT-IIA or KK questionnaire, we need to simply multiply the estimated weight $\hat{\mathbf{D}} \in \mathbb{R}^{N \times R}$ by the transpose of the component matrix found for the $[Score]$ domain, $\mathbf{E} \in \mathbb{R}^{T \times R}$.

$$\hat{\mathbf{S}} = \hat{\mathbf{D}} \times [\mathbf{E}]^T \tag{8}$$

where $\hat{\mathbf{S}} \in \mathbb{R}^{N \times T}$ is the estimated score for the new child(ren).

III. Results

A. Experiments

The above analysis was done on a combination of the 3 CCP blocks, such that analysis included evaluation on a cross-validated test using two CCP blocks and a repeated evaluation on the other withheld CCP block. In particular, a 10-fold cross-validation combining two CCP blocks of data withheld two or three subjects at each fold when training the JEDI model to establish the "within block" validation set. All data from the other CCP block was completely excluded from the model training as well, to provide the second validation test set designated as the "out-of-block" test. This double hold-out scheme promoted robustness in the results two ways. First, evaluating two repeated tests helps inform on the stability of the JEDI model across multiple evaluations. Second, by excluding a CCP block from the model training, potential systemic differences present in the CCP blocks may be revealed from comparing the within block and out-of-block validation results. Importantly, the children to be tested had their scores held out prior to the z-score transformation as well, to ensure no potential contamination between the training and testing sets. Figure 2 illustrates this experimental outline. The new estimated scores for both within the cross-validation and out-of-block testing were compared to the original true scores of each child. The joint CPD factorization was run on each training fold separately, and repeated using all permutations of the CCP blocks 1, 2, and 3.

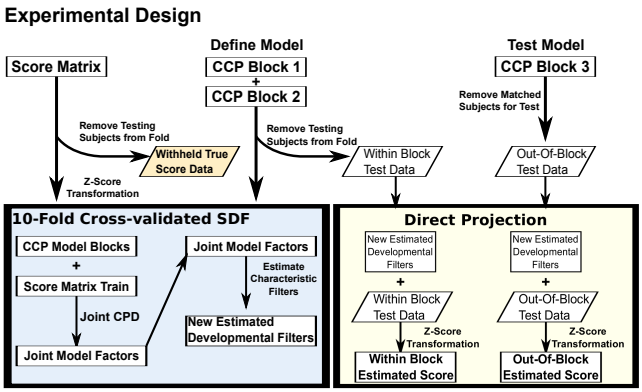

Fig. 2. A flowchart illustrating the experimental design of the JEDI training and testing. The true scores for comparison are highlighted in the tan parallelogram, showing they were not included for any steps of the JEDI model training. The blue box highlights the key steps in the JEDI joint factorization for each training fold. The yellow box highlights how withheld testing data was used to estimate scores for both the 10-fold cross-validation and the out-of-block test data.

B. Parameter Weights

The joint factorization reported for the JEDI model used $R = 4$ components, and hyperparameter weights of 3.162, 3.162, 0.1, and 1 for λ_{1-4}, respectively. These values were chosen based in part on results from the complete grid search. To maintain commonality across the various testing experiments, we opted to use a single set of weights which had been found within five lowest differences between real and estimated values for the decomposition and validation blocks (e.g. use CCP Block 1 for joint decomposition, CCP Block 2 for validation, CCP Block 3 for independent testing) in the grid search. This allowed direct comparisons of results across permutations, scores and testing blocks since each JEDI joint factorization weighted the profiles identically.

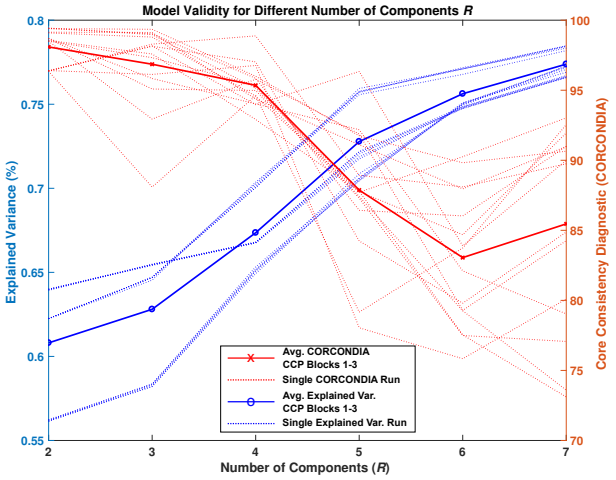

Fig. 3. Comparison between the percent explained variance and Core Consistency Diagnostic (CORCONDIA) of the tensor data in the JEDI model from varying the number of components R over a given set of weights λ_{1-4}. The group mean across a set of 5 repeated decompositions for each CCP Block 1, 2 and 3 are provided.

Figure 3 illustrates the EV and CORCONDIA for an optimal set of weights λ_{1-4} for various numbers of components using 5 repeated decompositions for each CCP block (15 repetitions total). The results from Figure 3 highlight that using greater than 5 components in the decomposition lead to diminishing returns in EV with large decreases in the CORCONDIA. As such, Figure 3 highlights why the grid search only included up to 5 components for analysis.

C. Developmental Score Prediction

Figure 4 illustrates the results for the predicted scores, based on the permutations of model training and testing from CCP blocks 1, 2, and 3 using the procedure outlined in (4)-(8). The three permutations of the JEDI model training and testing are represented in each row of Figure 4. The left hand column illustrates the predicted scores for using EEG from the 'within block' of the cross-validated data as outlined in Figure 2. The right column shows the predicted score results using EEG data from the separate 'out-of-block' testing fold, as seen in 2. For each panel, the actual and predicted scores for each child is displayed side-by-side for the WASI-II (left-most data), WIAT-IIA (central data), and KK questionnaire (right-most data) tests. Box-plots were used to capture overall trends in the real and estimated scores. A scatter-plot linking real-to-predicted data was overlaid on top the associated box-plot pairs in order to show the efficacy of the JEDI model at an individual level.

The majority of trends in Figure 4 indicate the JEDI could estimate a child's developmental score. For the WASI-II test, a median difference of ± 20.81, and ± 21.68 points on average across permutations for within and out-of-block tests, respectively. Similarly, for the WIAT-IIA test a median difference of ± 15.99, and ± 16.85 points across the different permutations for the within and out-of-block tests, respectively. The KK questionnaire had the lowest median difference in points at ± 8.89, and ± 11.07 across permutations for the within and out-of-block tests, respectively. The lowest median value for

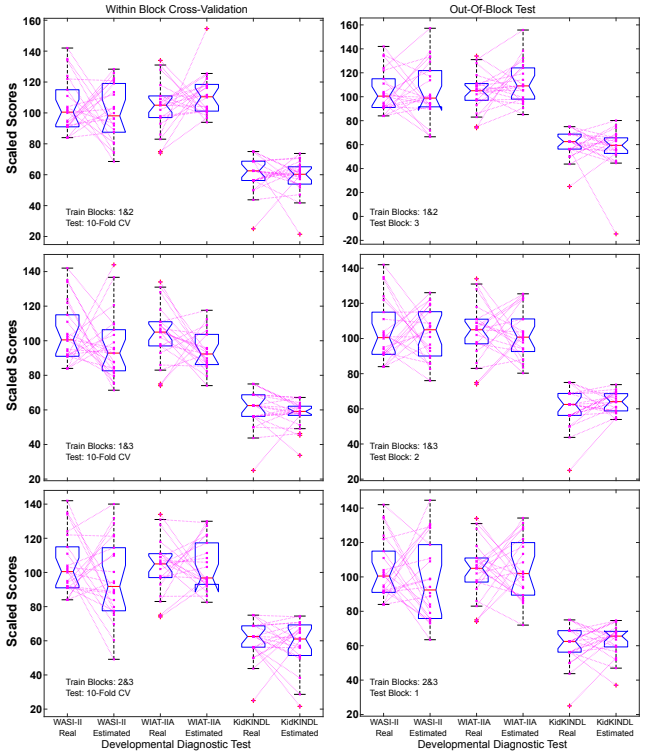

Fig. 4. Actual and predicted scores for all subjects for three clinical measures of development, for each training and testing permutation of CCP blocks for the JEDI model. Rows reflect the training/testing permutation. The left column gives results from the 10-fold within block cross-validation, and the right column for the out-of-block data. For each panel, from left to right are actual and predicted value pairs for WASI-II, WIAT-IIA and the KidKINDL questionnaire, respectively. Each pair of box-plots includes the subject distribution in magenta, with a dotted line indicating the child's 'actual-to-estimated' score values. No significant differences between the actual and predicted values were found, based on a two-tailed Student's t-test at a significance level of $p = 0.05$, corrected for multiple comparisons via False Discovery Rate (FDR, at $p = 0.05$)

the KK questionnaire was ± 4.79 (in the within block cross-validation test), while the highest was ± 13.91 in the out-of-block test. These results are summarized in Figure 5 as a complimentary graphic to the main results in Figure 4.

Across all model testing permutations, only 1 pair of real and estimated scores were found to be significantly different based on a two-tailed Student's t-test, at a level of $p = 0.05$. This was the WIAT-IIA scores for the within block at $p \approx 0.0180$, using CCP blocks 1, 3 for training and block 2 for testing. However, correcting for multiple comparisons via False Discovery Rate (FDR) or Bonferroni found no significant differences from the distribution of real scores to the JEDI estimated values.

Incorporating a child's age into the score matrix $\mathbf{S}$ was also examined as a measure of verifying the suitability of the JEDI model in estimating aspects of child development. Including child age as an alternative score in the matrix $\mathbf{S}$ was found to have no significant effect on score estimations made by the JEDI model. Further, the median estimated ages differed from the real ages by ± 1.36 and ± 1.12 on average across all permutations for the within block and out-of-block tests, respectively. These results are included in Supplementary

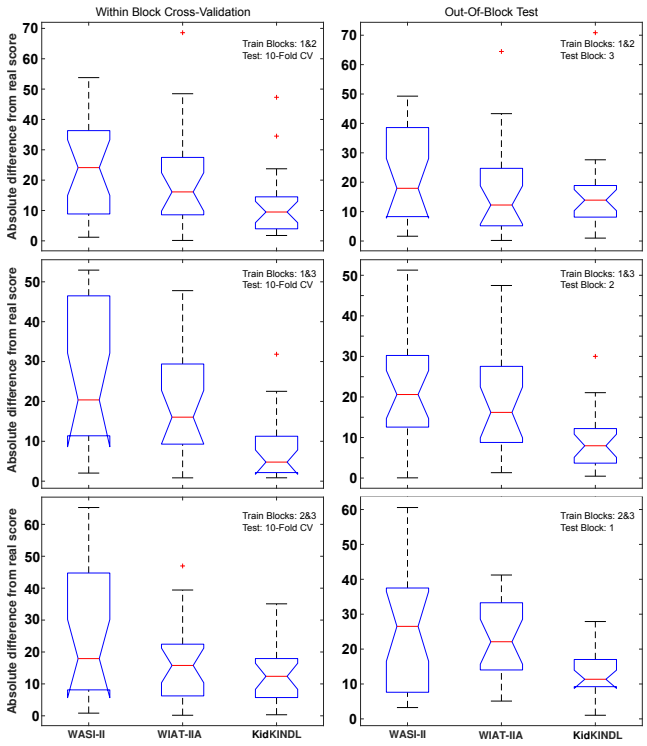

Fig. 5. Absolute difference between the real and estimated scores given by the JEDI model, for all permutations of CCP blocks. The three left-hand panels show results using the within block cross-validation, while the right-hand panel shows results using the out-of-block test data. For each panel, from left to right the scores are: WASI-II, WIAT-IIA and KidKINDL questionnaire.

Figures A and B.

IV. Discussion

This paper presents a new model built upon the joint factorization of EEG data and clinical measures of development with the aim to infer developmental status of children using EEG data. The Joint EEG-Development Inference (JEDI) model was described in detail, and tested on predicting WASI-II, WIAT-IIA and KK questionnaire scores rapidly via characteristic filters. The results demonstrate an important step towards using EEG data to predict classic developmental scores of children. The presented JEDI model provided a preliminary example for estimating a child's developmental score via EEG, with results found to be within an adjacent category (e.g. 'average','above average', etc.) of the real data. Furthermore, at a population level, the distribution of developmental scores predicted by the JEDI model were not significantly different from the original distribution of scores in the paediatric population.

A. JEDI Trends

Results from the JEDI model are an encouraging, initial step. Although the median values across the tests were modest with respect to predicting scores, there were some important trends in the data which promote further investigation. For example, across the WASI-II and WIAT-IIA tests, there appeared to be a trend for individual subjects in the interquartile range (i.e. the 'middle' 50% of data) to have lower differences between the real and estimated scores. In contrast, the more 'extreme' scores (either above or below) tended to be inversely estimated (e.g. low scores were estimated significantly higher by the JEDI model, while high scores were estimated significantly lower). Pearson's R reveals a negative correlation trend between the estimated and real scores in the WASI-II (within block testing $R = -0.3403 \pm 0.138$, and out-of-block testing $R = -0.1278 \pm 0.1694$) and WIAT-IIA (within block testing $R = -0.2639 \pm 0.2019$, and out-of-block testing $R = -0.1073 \pm 0.1425$) across all CCP block combinations. These trends are significant for the uncorrected $p - values$, but fail to remain significant when corrected for multiple comparisons for a FDR of $p = 0.05$. This helps explain, in part, why many of the large differences in individual estimates of developmental scores tended to come from data originally in the top or bottom 25% of the population distribution. This trend, however, is not apparent for the KK questionnaire, where estimated data appear to be more positively correlated at $R = 0.1411 \pm 0.2627$, and $R = 0.1805 \pm 0.1038$ for within and out-of-block testing respectively. Furthermore, the KK questionnaire data seem more likely to be in a similar range as the real data regardless of their positioning in the original data distribution. As noted below, there are future work possibilities which may help improve the JEDI model's predictive capabilities and consequently could improve these findings.

It is important to contextualize the estimated results and differences from the JEDI model with respect to a real-world interpretation. In practice, the median differences found between the real and estimated score pairs for WASI-II demonstrate a reasonable ability for the JEDI to retain a child's 'descriptive category' with respect to the original FSIQ-4 values. The median change of approximately ± 21 points for JEDI model estimates of WASI-II imply that children assigned an 'average' score can be categorized as 'below average', 'average', or even 'above average'. The normal curve distribution of WASI-II has a standard deviation of 15 points, putting the JEDI estimation outside the expected deviation. The consistency from both the within block cross-validation and out-of-block test data supports that, given the specific parameters for the presented JEDI model, these results are likely stable.

The WIAT-IIA predicted scores were consistently closer to their real counterparts via the JEDI model. Given that the standard deviation for WIAT-IIA is reported at 15 points, the JEDI model median differences fell critically within only 1-2 points of the expected standard deviation curve, evaluated for the within and out-of-block tests, respectively. This implies the shared developmental factors underlying the EEG were more readily captured via the JEDI joint factorization. This effect seemed to be constant for both within block cross-validation and for out-of-block testing, indicating it also was a stable estimation.

The estimated KidKINDL questionnaire values were also consistently within one categorical step of their respective real counterpart. The within block cross-validation and out-of-block testing median differences seem slightly less stable than the WASI-II or WIAT-IIA predictions (e.g. the approximate

4 point difference in median change). When scaled back to a 5 point scale the predictive capability seems comparable to the other tests, in that estimations will fall within the 1-2 category range (median change of ± 1.36, and ± 1.44 for within and out-of-block testing, respectively). This instability could be partial a result from the specific model joint factorization properties, which could change with different model selection (e.g. different weights of λ_{1-4}, R, etc.).

For all developmental tests, the outlier results represent a key situation where the model completely failed. However, from a clinical point of view, these outliers are such drastic failures, that they are easy to immediately identify. Extremely low developmental scores (seen in several predicted children in Figure 4) are clearly not normal, so a decision can be made to investigate with the classical approach of running the developmental tests. This model therefore could help to reduce the resources required for clinical developmental assessment.

B. Model Selection and Stability

Model selection is a critical aspect to consider when evaluating the given results. We opted to use a single set of hyperparameters across all combinations of data to retain the ability to more intuitively compare across outcomes. This allowed less complex interpretation when comparing data, at the trade-off of perhaps better fitting of the JEDI model to each training/testing fold and permutation, which in turn could improve the estimated scores. Another example, is at the tensor decomposition level. We selected the CPD model for its ability to exploit the characteristic filters through direct projection [8]. However, due to CPD's core design [8], [33], the model may be less flexible than desired. Using a more flexible tensor model, like PARAFAC2 [16], [39] or block-term decomposition [31], [40] offer directions for future work which may improve the estimations from the JEDI model.

Stability in the JEDI model is another key aspect to consider. Despite testing with two separate scenarios, including the completely unseen EEG data, the median results from the JEDI model remained relatively stable (i.e., similar results were found for both testing groups). The general stability of the model is further supported in that none of the estimated distributions were significantly different from the actual values, with distributions relatively similar in size and location in Figure 4. Importantly, however, is that the individual predictions can vary if the JEDI model is run repeatedly, due to the joint tensor factorization and random initialization of the score matrix. Also, since the SDF is a numerical solution estimating the common underlying factors, the joint tensor model solution is not guaranteed to be identical for each run. In this context, some of the outlier values predicted by the JEDI model may be a factor resulting from something akin to a local minima in the data factorization. A direction for future work could be to explore how individual subject estimations change given particular decompositions, and how to determine which is closest to the 'true' developmental value.

Additionally, the JEDI model's ability to effectively estimate the KK questionnaire, a self-reported measure, better than the WASI-II or WIAT-IIA IQ tests was a surprise. However, the (relative) instability of the JEDI's estimation for the KK questionnaire between the within block and out-of-block tests highlights that this phenomena may be a result of the relatively small sample size available for this pilot project. Evaluating the model on an external data source using 4 children who were not included in the experiments above (due to their inability to complete all 3 of the CCP blocks) revealed the WASI-II and WIAT-IIA were consistent with the results reported above, with the KK questionnaire estimations found to be more variable. Evaluating the JEDI model on a larger population should help clarify the nature of these findings.

C. Application and Relevance

The ability to estimate a child's developmental status from EEG would be beneficial for many applications. In terms of brain-computer interface (BCI) or other rehabilitation strategies which EEG is featured predominantly, the JEDI model could open up a means to rapidly assess how the intervention was progressing directly. This could help alleviate some of the resource intensive requirements for repeated clinical assessments. The present gold standard is to do labor intensive detailed psychometric tests, which are not always available or possible and may suffer from learning effects over repeat testing. The proposed model thus opens up the potential for a computational biomarker for developmental status clinically. Development of such a biomarker is a sought after, yet elusive item in trials aimed at assessing interventions to developmental problems or at looking at adverse effects on development in drug trials.

For EEG based rehabilitation paradigms, it offers an additional use for the recorded EEG data. The data could be used to both infer development and to record specific tasks of interest. Furthermore, the JEDI model provides an avenue to more directly link changes found at an electrophysiological level to well-established measurements of development at the psychological level. The proposed JEDI model thus offers multiple benefits in its data-driven approach to inferring a child's developmental status using a common EEG task.

D. Limitations and Future Work

There were several limitations associated with this work. First, the sample size of this study was relatively small ($n = 22$). Due to the exclusive design of the study, over half of the available children were not included in analysis since they were missing data from either a CCP block, a developmental score or both. Future work should repeat these findings in a larger group to verify the results across a larger population. Additionally, the sample investigated was for children with healthy development, and an older childhood population. As development and EEG are heavily linked to age of the child, it is unclear if the JEDI model would be as predictive in children with neurological deficits (e.g. epilepsy or paediatric stroke), or in children who are very young (< 5 years old). Additional investigations should be conducted to demonstrate the model in a variety of paediatric populations.

Another limitation was that the grid search was done prior to the actual analysis, in order to identify a common

hyperparameter set. Although computationally expensive, an alternative could be to implement the wide grid search for optimizing hyperparameters at each fold of the analysis. The current work was intended to demonstrate broadly the potential effectiveness of the JEDI model, and as such future work could implement this increased complexity by having the SDF training and validation held within each cross-validation fold and hyperparameters changed accordingly. Furthermore, the grid search could be expanded upon to include additional properties for optimization such as changing the frequency range explored or which channels are used in analysis.

The similarity between the repeated within and out-of-block tests investigated in this work is another potential limitation. As all of the CCP block recordings were done within the same day, the current results are limited in reporting on longitudinal information or potential errors coming from intra-session variabilities. As such, including in a mixed-effect regression model to account for such potential difference in longitudinal recordings could be more thoroughly explored in follow-up studies.

Finally, the reported results focus only on the model's ability to predict developmental scores. Investigations on how to utilize the model to also predict a child's decision between left and right buttons could help bolster the relevance of the data.

V. Conclusion

This manuscript outlines using joint-factorization of multi-relational data reflecting child development in a model termed the Joint EEG-Development Inference (JEDI) model. The structure of the JEDI model was comprised of the fusion of a common EEG task (button press) and a matrix of clinically relevant developmental diagnostic tests (WASI-II, WIAT-IIA, KidKINDL). Through using multi-way tensors to maintain the inherent structural relationships present in the EEG data, shared developmental information could be gleaned from the SDF. Considerations were given to the relative weights for the tensor-matrix fusion and regularization factors (L2, L1), through a grid search approach. Results indicated that the JEDI model is a prominent first step towards inferring the developmental scores of children via EEG, with estimated values falling approximately within the same or adjacent categorical group as the real values. Analysis showed the JEDI model was stable across repeated recordings, as similar results were found when using 10-fold cross-validated data and unseen EEG data. The JEDI model represents the foundation for a promising tool which could be used in estimating developmental status of children rapidly, using data from a common EEG task.

Acknowledgment

The authors would like to thank the Child Mind Institute for providing the EEG and behavioural data, and the children and families who participated in the study. Code used in this manuscript's analysis is openly available upon request.

References

[1] P. R. Penn, F. D. Rose, and D. a. Johnson, "Virtual enriched environments in paediatric neuropsychological rehabilitation following traumatic brain injury: Feasibility, benefits and challenges." *Dev. Neurorehabil.*, vol. 12, no. 1, pp. 32–43, 2009.

[2] D. A. Rohani, H. B. D. Sorensen, and S. Puthusserypady, "Brain-computer interface using P300 and virtual reality: A gaming approach for treating ADHD," in *2014 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2014. IEEE, aug 2014, pp. 3606–3609.

[3] N. Mrachacz-Kersting, N. Jiang, A. J. T. Stevenson, I. K. Niazi, V. Kostic, A. Pavlovic, S. Radovanovic, M. Djuric-Jovicic, F. Agosta, K. Dremstrup, and D. Farina, "Efficient neuroplasticity induction in chronic stroke patients by an associative brain-computer interface," *J. Neurophysiol.*, vol. 115, no. 3, pp. 1410–1421, mar 2016.

[4] L. van Dokkum, T. Ward, and I. Laffont, "Brain computer interfaces for neurorehabilitation its current status as a rehabilitation strategy post-stroke," *Ann. Phys. Rehabil. Med.*, vol. 58, no. 1, pp. 3–8, feb 2015.

[5] E. Kinney-Lang, B. Auyeung, and J. Escudero, "Expanding the (kaleido)scope: exploring current literature trends for translating electroencephalography based brain-computer interfaces for motor rehabilitation in children," *J. Neural Eng.*, vol. 13, no. 6, p. 061002, 2016.

[6] M. Grosse-Wentrup, D. Mattia, and K. Oweiss, "Using brain-computer interfaces to induce neural plasticity and restore function." *J. Neural Eng.*, vol. 8, no. 2, p. 025004, 2011.

[7] E. Kinney-Lang, L. Spyrou, A. Ebied, R. Chin, and J. Escudero, "Elucidating age-specific patterns from background electroencephalogram pediatric datasets via PARAFAC," in *2017 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* IEEE, jul 2017, pp. 3797–3800.

[8] E. Kinney-Lang, L. Spyrou, A. Ebied, R. F. Chin, and J. Escudero, "Tensor-driven extraction of developmental features from varying paediatric EEG datasets," *J. Neural Eng.*, may 2018.

[9] E. Mikołajewska and D. Mikołajewski, "The prospects of brain computer interface applications in children," *Cent. Eur. J. Med.*, vol. 9, no. 1, pp. 74–79, 2013.

[10] L. Sorber, M. Van Barel, and L. De Lathauwer, "Structured data fusion," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 586–600, 2015.

[11] E. Acar, T. G. Kolda, and D. M. Dunlavy, "All-at-once Optimization for Coupled Matrix and Tensor Factorizations," *Mlg'11*, no. 1, may 2011.

[12] E. Acar, M. A. Rasmussen, F. Savorani, T. Næs, and R. Bro, "Understanding data fusion within the framework of coupled matrix and tensor factorizations," *Chemom. Intell. Lab. Syst.*, vol. 129, pp. 53–63, 2013.

[13] E. Acar, E. E. Papalexakis, G. Gürdeniz, M. a. Rasmussen, A. J. Lawaetz, M. Nilsson, and R. Bro, "Structure-revealing data fusion," *BMC Bioinformatics*, vol. 15, no. 1, p. 239, 2014.

[14] E. Acar, R. Bro, and A. K. Smilde, "Data Fusion in Metabolomics Using Coupled Matrix and Tensor Factorizations," *Proc. IEEE*, vol. 103, no. 9, pp. 1602–1620, 2015.

[15] N. Vervliet, O. Debals, and L. De Lathauwer, "Tensorlab 3.0 - Numerical optimization strategies for large-scale constrained and coupled matrix/tensor factorization," in *Conf. Rec. - Asilomar Conf. Signals, Syst. Comput.* IEEE, nov 2017, pp. 1733–1738.

[16] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2008.

[17] N. Langer, E. J. Ho, L. M. Alexander, H. Y. Xu, R. K. Jozanovic, S. Henin, A. Petroni, S. Cohen, E. T. Marcelle, L. C. Parra, M. P. Milham, and S. P. Kelly, "A resource for assessing information processing in the developing brain using EEG and eye tracking," *Sci. Data*, vol. 4, p. 170040, apr 2017.

[18] D. Wechsler, *Wechsler Abbreviated Scale of Intelligence - Second Edition (WASI-II)*. San Antonio, TX: NCS Pearson, 2011.

[19] ——, *Wechsler Individual Achievement Test - Second Edition Abbreviated (WIAT-IIA)*, 2nd ed. San Antonio, TX: Psychological Corporation., 2005.

[20] U. Ravens-Sieberer and M. Bullinger, "Assessing health-related quality of life in chronically ill children with the German KINDL: First psychometric and content analytical results," *Qual. Life Res.*, vol. 7, no. 5, pp. 399–407, jul 1998.

[21] ——, "News from the kindl-questionnaire: A new version for adolescents," *Quality of Life Research*, pp. 653–653, 1998.

[22] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data," *Comput. Intell. Neurosci.*, vol. 2011, pp. 1–9, 2011.

[23] M. Matsuura, K. Yamamoto, H. Fukuzawa, Y. Okubo, H. Uesugi, M. Moriiwa, T. Kojima, and Y. Shimazono, "Age development and sex differences of various EEG elements in healthy children and adults–quantification by a computerized wave form recognition method," *Electroencephalogr Clin Neurophysiol*, vol. 60, no. 5, pp. 394–406, 1985.
[24] S. J. Segalowitz, D. L. Santesso, and M. K. Jetha, "Electrophysiological changes during adolescence: A review," *Brain Cogn.*, vol. 72, no. 1, pp. 86–100, 2010.
[25] M. J. Taylor and T. Baldeweg, "Application of EEG, ERP and intracranial recordings to the investigation of cognitive functions in children," *Dev. Sci.*, vol. 5, no. 3, pp. 318–334, 2002.
[26] M. M. Pangelinan, F. A. Kagerer, B. Momen, B. D. Hatfield, and J. E. Clark, "Electrocortical Dynamics Reflect Age-Related Differences in Movement Kinematics among Children and Adults," *Cereb. Cortex*, vol. 21, no. 4, pp. 737–747, apr 2011.
[27] S. Bender, M. Weisbrod, H. Bornfleth, F. Resch, and R. Oelkers-Ax, "How do children prepare to react? Imaging maturation of motor preparation and stimulus anticipation by late contingent negative variation," *Neuroimage*, vol. 27, no. 4, pp. 737–752, 2005.
[28] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab 3.0," 2016.
[29] N. Vervliet, O. Debals, L. Sorber, and L. De Lathauwer, "Breaking the Curse of Dimensionality Using Decompositions of Incomplete Tensors: Tensor-based scientific computing in big data analysis," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 71–79, sep 2014.
[30] F. Cong, Q. H. Lin, L. D. Kuang, X. F. Gong, P. Astikainen, and T. Ristaniemi, "Tensor decomposition of EEG signals: A brief review," *J. Neurosci. Methods*, vol. 248, pp. 59–69, 2015.
[31] A. Cichocki, "Tensor Decompositions: A New Concept in Brain Data Analysis?" *arXiv Prepr. arXiv1305.0395*, vol. 50, no. 2011, pp. 507–517, may 2013.
[32] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, 2015.
[33] R. a. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an explanatory multimodal factor analysis," *UCLA Work. Pap. Phonetics*, vol. 16, no. 10, pp. 1– 84, 1970.
[34] N. Klaas, M. Faber, R. Bro, and P. K. Hopke, "Recent developments in CANDECOMP / PARAFAC algorithms : a critical review," *Chemom. Intell. Lab. Syst.*, vol. 65, pp. 119 – 137, 2003.
[35] A. Cichocki, Y. Washizawa, T. Rutkowski, H. Bakardjian, Anh-Huy Phan, Seungjin Choi, Hyekyoung Lee, Qibin Zhao, Liqing Zhang, and Yuanqing Li, "Noninvasive BCIs: Multiway Signal-Processing Array Decompositions," *Computer (Long. Beach. Calif).*, vol. 41, no. 10, pp. 34–42, oct 2008.
[36] R. Bro and U. V. Amsterdam, "Multi-way Analysis in the Food Industry," *Food Technol.*, no. november, 1998.
[37] G. Tomasi and R. Bro, "A comparison of algorithms for fitting the PARAFAC model," *Comput. Stat. Data Anal.*, vol. 50, no. 7, pp. 1700–1734, 2006.
[38] A. Cichocki, "Era of Big Data Processing: A New Approach via Tensor Networks and Tensor Decompositions," *arXiv*, pp. 1–30, mar 2014.
[39] J. Escudero, E. Acar, A. Fernández, and R. Bro, "Multiscale entropy analysis of resting-state magnetoencephalogram with tensor factorisations in Alzheimer's disease," *Brain Res. Bull.*, vol. 119, pp. 136–144, oct 2015.
[40] G. Zhou, A. Cichocki, Y. Zhang, and D. Mandic, "Group Component Analysis for Multi-block Data: Common and Individual Feature Extraction," vol. 27, no. 11, pp. 2426–2439, 2012.

Supplementary Matrial for: *Introducing the Joint EEG-Development Inference (JEDI) Model: A Multi-way, Data Fusion Approach for Estimating Paediatric Developmental Scores Via EEG*

Eli Kinney-Lang, Ahmed Ebied, Bonnie Auyeung, Richard F.M. Chin, and Javier Escudero *Member, IEEE*

SUPPLEMENTARY INFORMATION

Below are Supplementary Figures A and B, complementary to the main article.

E. Kinney-Lang, A. Ebied and J. Escudero are with the School of Engineering, Institute for Digital Communications (IDCOM), University of Edinburgh, Edinburgh, United Kingdom, EH9 3FG. Bonnie Auyeung is with the School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom EH8 9JZ. Richard Chin is affiliated with the Royal Hospital for Sick Children, Edinburgh, the Muir Maxwell Epilepsy Centre and the Centre for Clinical Brain Sciences at the University of Edinburgh, Edinburgh.

Primary author e-mail: e.kinney-lang@ed.ac.uk

Manuscript submitted June 30, 2018.

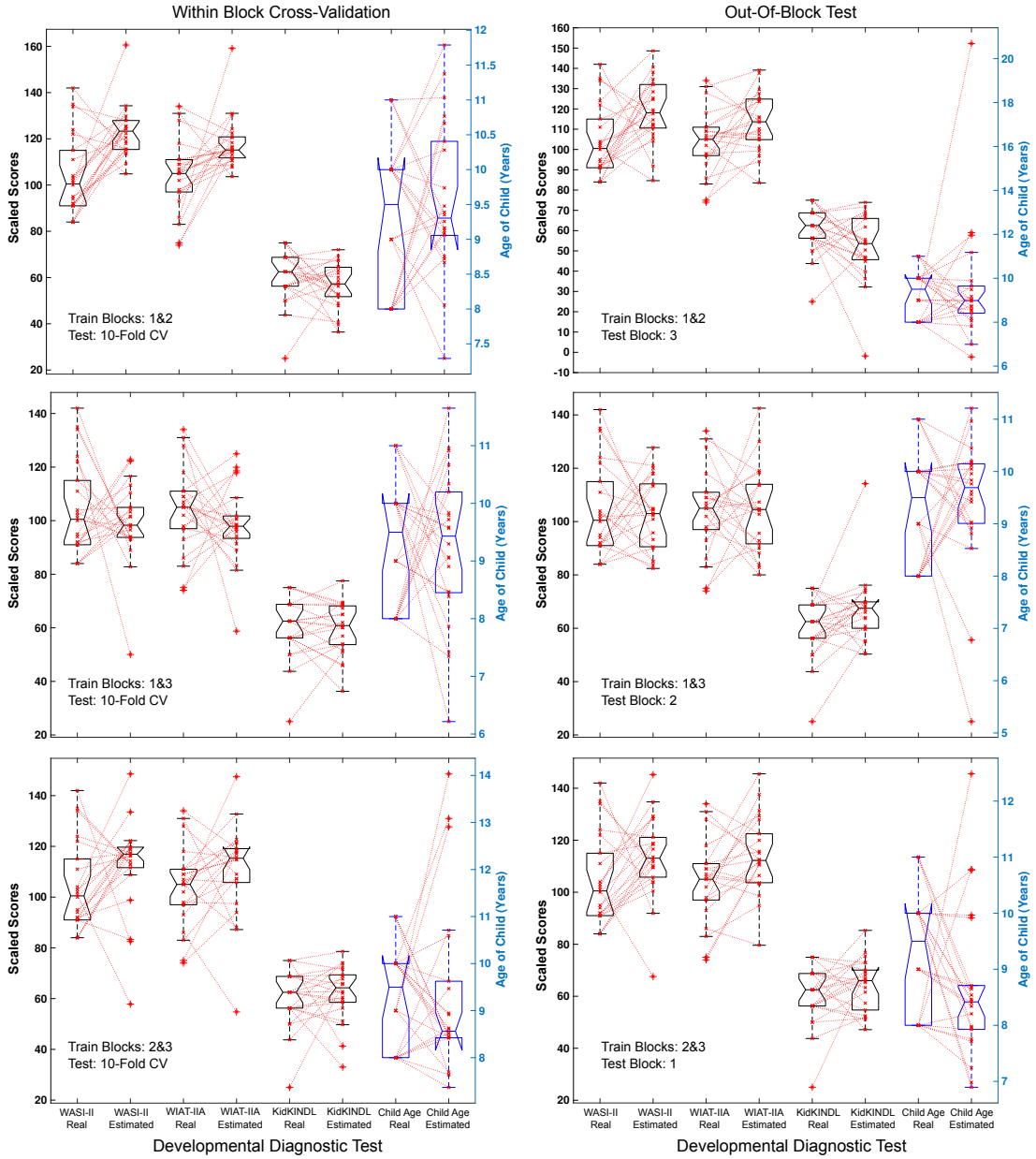

Fig. 1. Supplementary Figure A. Actual and predicted scores for all subjects for the three clinical measures of development and child age, across all CCP block permutations. Rows reflect the training/testing permutation. The left column gives results from the 10-fold within block cross-validation, and the right column for the separate out-of-block data. For each panel, from left to right are the actual and predicted value pairs for WASI-II, WIAT-IIA, KidKINDL questionnaire, and child age (blue) respectively. Each pair of box-plots includes the subject distribution in red, with a dotted line indicating the child's 'actual-to-estimated' score values. No significant differences between the actual and predicted values were found, based on a two-tailed Student's t-test at a significance level of $p = 0.05$, corrected for multiple comparisons via False Discovery Rate (FDR, at $p = 0.05$)

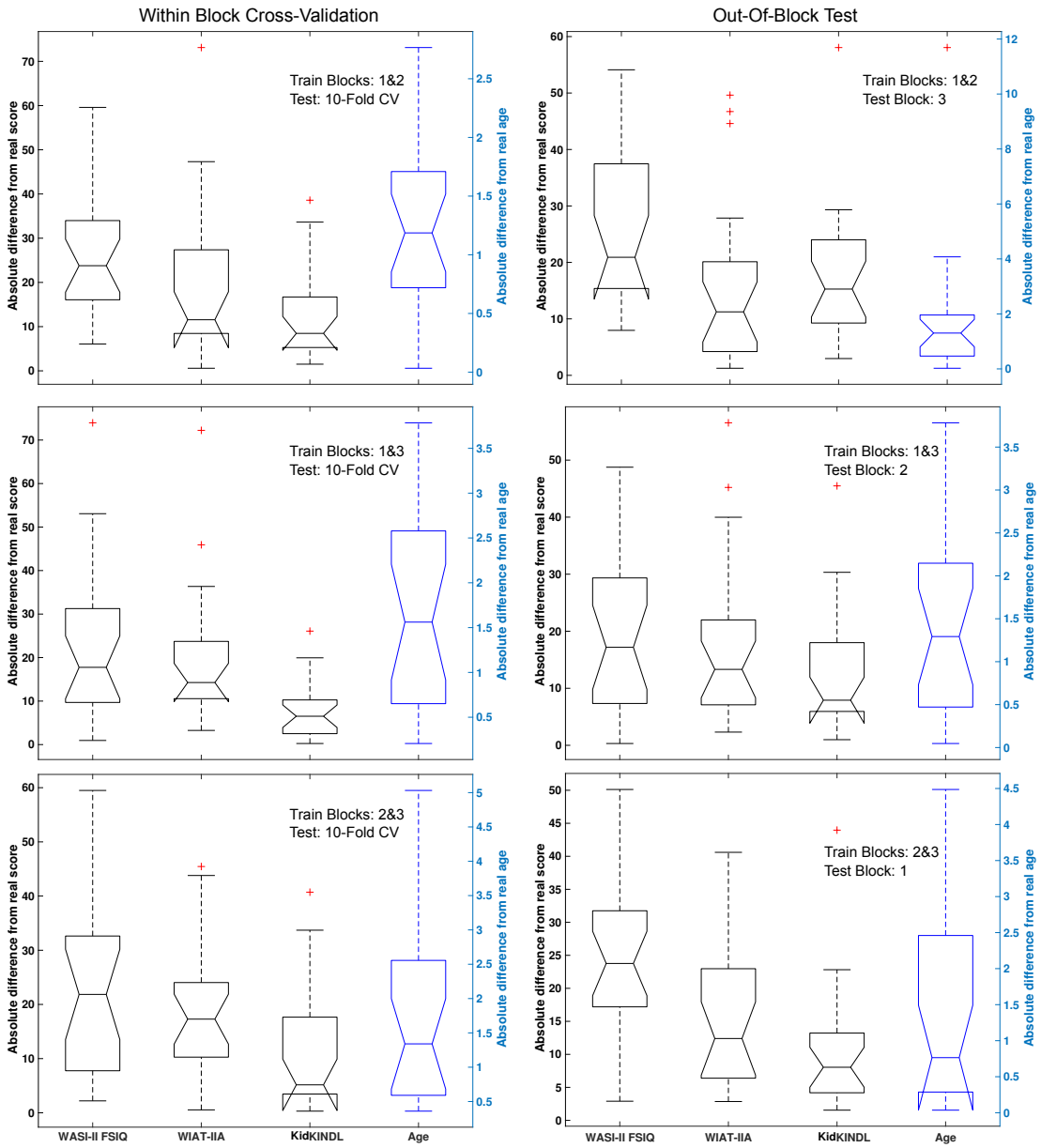

Fig. 2. Supplementary Figure B. Absolute difference between the real and estimated scores and child age given by the JEDI model, for all permutations of CCP blocks. The three left-hand panels show results using the within block cross-validation, while the right-hand panel shows results using the out-of-block test data. For each panel, from left to right the scores are: WASI-II, WIAT-IIA, KidKINDL questionnaire and Age.